



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2012

---

## **Optimal weighting in galaxy surveys: Application to redshift-space distortions**

Hamaus, Nico ; Seljak, Uroš ; Desjacques, Vincent

**Abstract:** Using multiple tracers of large-scale structure allows to evade the limitations imposed by sampling variance for some parameters of interest in cosmology. We demonstrate the optimal way of carrying out a multitracers analysis in a galaxy redshift survey by considering the principal components of the shot noise matrix from two-point clustering statistics. We show how to construct two tracers that maximize the benefits of sampling variance and shot noise cancellation using optimal weights. On the basis of high-resolution N-body simulations of dark matter halos we apply this technique to the analysis of redshift-space distortions and demonstrate how constraints on the growth rate of structure formation can be substantially improved. The primary limitations are nonlinear effects, which cause significant biases in the method already at scales of  $k < 0.1 h \text{Mpc}^{-1}$ , suggesting the need to develop nonlinear models of redshift-space distortions in order to extract the maximum information from future redshift surveys. Nonetheless we find gains of a factor of a few in constraints on the growth rate achievable when merely the linear regime of a galaxy survey like EUCLID is considered.

DOI: <https://doi.org/10.1103/PhysRevD.86.103513>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-70068>

Journal Article

Originally published at:

Hamaus, Nico; Seljak, Uroš; Desjacques, Vincent (2012). Optimal weighting in galaxy surveys: Application to redshift-space distortions. *Physical Review D (Particles, Fields, Gravitation and Cosmology)*, 86:103513.

DOI: <https://doi.org/10.1103/PhysRevD.86.103513>

# Optimal Weighting in Galaxy Surveys: Application to Redshift-Space Distortions

Nico Hamaus,<sup>1,\*</sup> Uroš Seljak,<sup>1,2,3</sup> and Vincent Desjacques<sup>4</sup>

<sup>1</sup>*Institute for Theoretical Physics, University of Zurich, 8057 Zurich, Switzerland*

<sup>2</sup>*Physics Department, Astronomy Department and Lawrence Berkeley National Laboratory,  
University of California, Berkeley, California 94720, USA*

<sup>3</sup>*Ewha University, Seoul 120-750, S. Korea*

<sup>4</sup>*Département de Physique Théorique & Center for Astroparticle Physics,  
Université de Genève, 24 Quai Ernest Ansermet, 1211 Genève 4, Switzerland*

(Dated: November 22, 2012)

Using multiple tracers of large-scale structure allows to evade the limitations imposed by sampling variance for some parameters of interest in cosmology. We demonstrate the optimal way of carrying out a multitracer analysis in a galaxy redshift survey by considering the principal components of the shot noise matrix from two-point clustering statistics. We show how to construct two tracers that maximize the benefits of sampling variance and shot noise cancellation using optimal weights. On the basis of high-resolution  $N$ -body simulations of dark matter halos we apply this technique to the analysis of redshift-space distortions and demonstrate how constraints on the growth rate of structure formation can be substantially improved. The primary limitations are nonlinear effects, which cause significant biases in the method already at scales of  $k < 0.1 h\text{Mpc}^{-1}$ , suggesting the need to develop nonlinear models of redshift-space distortions in order to extract the maximum information from future redshift surveys. Nonetheless we find gains of a factor of a few in constraints on the growth rate achievable when merely the linear regime of a galaxy survey like EUCLID is considered.

PACS numbers: 98.80.-k, 98.62.-g, 98.65.-r

## I. INTRODUCTION

One of the deepest mysteries of contemporary cosmology is the nature of the observed accelerated expansion of the Universe. So far, its evolution can be described remarkably well by Einstein's theory of gravitation including a nonzero cosmological constant  $\Lambda$ . However, in order to fit the astronomical observations (e.g., [1]),  $\Lambda$  must be many orders of magnitude smaller than what our standard model of particle physics would expect. This *hierarchy problem* inspired various departures from the cosmological standard model, such as modifications of Einstein's field equations or the introduction of exotic forms of matter.

A particularly sensitive probe of cosmology is the large-scale structure (LSS) of the Universe. Galaxy redshift surveys map out large fractions of its observable volume and thereby reconstruct a three-dimensional map of density fluctuations whose statistical properties directly relate to fundamental cosmological parameters [2–4]. Unfortunately, this reconstruction is hampered by the fact that galaxies are *biased* and *stochastic* tracers of the dominating dark matter density field. Even if the density field could be inferred perfectly well, the finite number  $N_k$  of independent Fourier modes in the survey sets a fundamental lower limit on the achievable uncertainty,

which is known as *sampling variance* (or *cosmic variance*), in case the survey size is the whole observable Universe). For example, in a measurement of the dark matter power spectrum  $P$ , the sampling variance limit is given by  $\sigma_P/P \geq \sqrt{2/N_k}$ , an uncertainty floor that propagates into all the parameters one wants to infer from  $P$ . This limit decreases towards smaller scales as more and more Fourier modes can be sampled, but at the same time linear theory starts to break down and higher-order perturbation theory has to be adopted to model  $P$  (see, e.g., [5–8]). Further complication arises in relating the observed galaxy power spectrum to the latter, as galaxy bias becomes nonlinear and nonlocal [9, 10].

An alternative approach to accurately probe cosmological parameters is to consider multiple tracers of the density field within the well-understood linear regime. The relative clustering amplitude between multiple tracers can be inferred without sampling variance limitation because the underlying density fluctuations cancel out in taking ratios [11, 12]. Therefore, any cosmological information that remains in the relative clustering amplitude between different tracers can potentially be inferred with a much higher accuracy.

In this paper we focus on a particular contribution to the clustering amplitude of galaxies coming from redshift-space distortions (RSD). These are caused by peculiar velocities along the line of sight, causing their clustering statistics to become anisotropic. First treated as a contamination, this effect has been realized to be a powerful probe of cosmology, as an understanding of RSD allows

---

\* hamaus@physik.uzh.ch

to infer the growth rate of structure formation, which is directly tied to the expansion history of the Universe as well as the theory of gravity (e.g., [13–23]).

After recapping the fundamentals of RSD and introducing a general formalism for the multitracers analysis in Sec. II, we present our results on the RSD analysis from  $N$ -body simulations in Sec. III. Finally we draw our conclusions in Sec. IV.

## II. FORMALISM

### A. Galaxies in redshift space

In galaxy surveys, radial distances are inferred via the individual redshift of objects, assuming they follow the Hubble flow. Due to gravitational attraction, however, galaxies (respectively, their host halos) build up peculiar velocities  $\mathbf{v}$  which contribute to their redshift via the Doppler effect. Hence, the real-space and redshift-space locations  $\mathbf{r}$  and  $\mathbf{s}$  of a galaxy are related as

$$\mathbf{s} = \mathbf{r} + \frac{\mathbf{v} \cdot \hat{\mathbf{r}}}{H(z)} \hat{\mathbf{r}}, \quad (1)$$

where  $\hat{\mathbf{r}}$  is the unit vector along the line of sight and  $H(z)$  is the Hubble constant as a function of redshift  $z$ . On large scales, gravity causes coherent infall of test particles into the potential wells of the dark matter. Thus, galaxies are moving towards overdense regions in the Universe, resulting in an enhancement of their inferred overdensity along the line of sight. According to linear perturbation theory, the redshift-space and real-space galaxy overdensities are related as

$$\delta_g^{(s)}(\mathbf{k}, \mu) = \delta_g^{(r)}(\mathbf{k}) + f\mu^2 \delta(\mathbf{k}), \quad (2)$$

where  $\mu$  the cosine of the angle between any wave vector  $\mathbf{k}$  in the survey and the line of sight,  $\delta$  the dark matter overdensity field and  $f$  the *growth rate* of structure. This well-known result [24] further makes use of the plane-parallel approximation, assuming the separation of any galaxy pair to be much smaller than their distance to the observer. On nonlinear scales, random motions are generated in the process of virialization, which causes a damping in the clustering amplitude of galaxies along the line of sight. This so-called Finger-of-God effect is often modeled phenomenologically by an additional Gaussian damping factor in Eq. (2) [25], but more elaborate schemes have been developed (e.g., [26–34]). We will neglect nonlinear corrections in this paper and focus on the large linear scales.

The growth rate  $f$  is the logarithmic derivative of the growth factor  $D$  with respect to the scale factor  $a$ . In linear theory, it can be expressed as

$$f \equiv d \ln D / d \ln a \simeq \Omega_m^\gamma, \quad (3)$$

with the matter density parameter  $\Omega_m$  and the *growth index*  $\gamma$  [35]. In Einstein gravity the value of  $\gamma$  is about 0.55, but can take on distinctly different values in modified gravity scenarios [36]. Therefore, constraints on the growth rate can provide viable tests on the theory of gravitation.

Galaxies only form in specific, discrete locations of the density field; they are referred to as *biased* and *stochastic* tracers of the dark matter. On linear scales, this relation can be described locally by

$$\delta_g^{(r)}(\mathbf{r}) = b_g \delta(\mathbf{r}) + \epsilon, \quad (4)$$

where the factor  $b_g$  is the *linear galaxy bias* and  $\epsilon$  a random variable denoted as *shot noise*, describing the stochastic nature of this relation. Both  $b_g$  and  $\epsilon$  depend on redshift, as well as various properties of the type of galaxies one is considering (e.g., luminosity, color, host-halo mass).

Together with Eq. (2) this yields a model for the overdensity field of galaxies in redshift space,

$$\delta_g^{(s)}(\mathbf{k}, \mu) = (b_g + f\mu^2) \delta(\mathbf{k}) + \epsilon. \quad (5)$$

Since the phenomena of galaxy biasing and RSD are to multiply the density field  $\delta$  by some factor, we can simply define a more general *effective bias* parameter  $b$  that contains both contributions,

$$b \equiv b_g + f\mu^2. \quad (6)$$

In the following, we will drop the superscript that distinguishes between real-space and redshift-space quantities for clarity. If not explicitly mentioned otherwise, all symbols should be understood as given in redshift space.

### B. Multiple tracers

In order to exploit the gains of sampling variance and shot noise cancellation, we need to consider multiple tracers of the dark matter density field [11, 12]. One way to achieve this is splitting some galaxy catalog into bins of a certain observable property of the galaxies (like luminosity, color, host-halo mass, etc.). However, the following framework is not limited to galaxies and may be adopted for other tracers of the dark matter density field as well.

#### 1. Covariance matrix

We start by writing the density fields of  $N$  tracers as a vector  $\boldsymbol{\delta}_g \equiv (\delta_{g_1}, \delta_{g_2}, \dots, \delta_{g_N})$ . The outer product of this vector, once ensemble averaged within a  $k$ -shell in Fourier space, yields the covariance matrix  $\mathbf{C} \equiv \langle \boldsymbol{\delta}_g \boldsymbol{\delta}_g^\dagger \rangle$ , where the  $^\dagger$  symbol denotes the Hermitian conjugate (transpose and complex conjugation). Plugging in the model for

galaxy overdensities in redshift space from Eq. (5), it reads

$$\mathbf{C} = \mathbf{b}\mathbf{b}^\dagger P + \mathcal{E}, \quad (7)$$

with the effective bias vector  $\mathbf{b} = \mathbf{b}_g + f\mu^2\mathbf{I}$ , the dark matter power spectrum  $P \equiv \langle \delta\delta^* \rangle$  and the *shot noise matrix*  $\mathcal{E} \equiv \langle \epsilon\epsilon^\dagger \rangle$  (by definition  $\langle \epsilon\delta^* \rangle = 0$ ). The covariance matrix contains all auto-power and cross-power spectra of the considered tracers, so in total  $N(N-1)/2$  independent elements per  $k$ -shell. However, since all tracers follow the same dark matter density distribution, these elements are correlated.

Let us compare the analysis for a single tracer with the one for two tracers. From a single tracer we can only observe an estimator of its auto-power spectrum

$$C(k, \mu) = (1 + \beta\mu^2)^2 b_g^2 P(k) + \mathcal{E}, \quad (8)$$

where we define  $\beta \equiv f/b_g$ . In this parametrization it is obvious to see that  $b_g$  is degenerate with the power spectrum  $P$ ; respectively, its normalization  $\sigma_8$  defined via

$$P \equiv \sigma_8^2 P_0, \quad (9)$$

where  $P_0$  describes the shape of the power spectrum and is assumed to be known, at least up to linear order. From Eq. (8), only the combination  $b_g^2 P$  can be determined at  $\mu = 0$  (usually, the shot noise is assumed to be Poissonian, meaning it is scale independent and given by the inverse number density of galaxies  $\bar{n}^{-1}$ ). The same is true for the growth rate  $f$ , we can only determine the product

$$f^2 P = \frac{C(k, \mu) - \mathcal{E} - [C(k, 0) - \mathcal{E}](1 + 2\beta\mu^2)}{\mu^4}, \quad (10)$$

but this does not apply for  $\beta$ , which can be extracted directly from observations of  $C$  where  $\mu \neq 0$  (e.g., [37]). Note, however, that in this case the achievable error on  $\beta$  is limited by the sampling variance inherent to  $P$ .

In case two distinct tracers of the density field with biases  $b_g$  and  $\alpha b_g$  are observed, where  $\alpha$  is their relative galaxy bias, we obtain the three following power spectra:

$$C_{11}(k, \mu) = (1 + \beta\mu^2)^2 b_g^2 P(k) + \mathcal{E}_{11}, \quad (11)$$

$$C_{22}(k, \mu) = (\alpha + \beta\mu^2)^2 b_g^2 P(k) + \mathcal{E}_{22}, \quad (12)$$

$$C_{12}(k, \mu) = (1 + \beta\mu^2)(\alpha + \beta\mu^2) b_g^2 P(k) + \mathcal{E}_{12}. \quad (13)$$

The degree to how well they are correlated is quantified by the cross-correlation coefficient  $r^2 \equiv C_{12}^2 / C_{11}C_{22}$ , so in the idealistic case of no shot noise ( $\mathcal{E} = 0$ ),  $r = 1$  and the ratios

$$C_{12}/C_{11} = C_{22}/C_{12} = \sqrt{C_{22}/C_{11}} = \frac{\alpha + \beta\mu^2}{1 + \beta\mu^2} \quad (14)$$

all yield the same expression, which is independent of  $P$ . Hence, a combination of observations at different  $\mu$  yields  $\alpha$  and  $\beta$  without sampling variance [12]. Unfortunately, in realistic surveys  $\mathcal{E} \neq 0$ , so  $P$  will not cancel out completely and the three ratios will be different and scale dependent. This gives rise to residual sampling variance inherent to a measurement of  $\alpha$  and  $\beta$ , which in general can be much smaller than in the single-tracer case. In turn, a better measurement of  $\beta$  allows a more precise estimate on  $f^2 P$  [see Eq. (10)]. However, the accuracy on  $f^2 P$  is still limited by sampling variance, yielding  $\sigma_{f^2 P} / f^2 P \geq \sqrt{2/N_k}$ , respectively  $\sigma_{f\sigma_8} / f\sigma_8 \geq \sqrt{1/2N_k}$  [12].

If both  $\delta_g$  and  $\delta$  are known, we can simply add the dark matter overdensity mode  $\delta$  to the overdensities of the tracers and write  $\boldsymbol{\delta} \equiv (\delta, \delta_g)$ . In this case, the effective bias can be obtained directly by taking the ratio  $\langle \delta_g \delta^* \rangle / \langle \delta \delta^* \rangle$ . The covariance matrix then becomes

$$\mathbf{C} = \begin{pmatrix} \langle \delta \delta^* \rangle & \langle \delta \delta_g^\dagger \rangle \\ \langle \delta_g \delta^* \rangle & \langle \delta_g \delta_g^\dagger \rangle \end{pmatrix} = \begin{pmatrix} P & \mathbf{b}^\dagger P \\ \mathbf{b} P & \mathbf{b} \mathbf{b}^\dagger P + \mathcal{E} \end{pmatrix}, \quad (15)$$

with  $N(N+1)/2$  independent elements. Now the degeneracy between  $b_g$ ,  $\sigma_8$  and  $f$  is lifted because  $P$  is known separately. Considering the cross-correlation coefficient of a tracer  $\delta_g$  with the dark matter  $\delta$ , we find

$$r^2 = \frac{\langle \delta_g \delta^* \rangle^2}{\langle \delta_g \delta_g^* \rangle \langle \delta \delta^* \rangle} = \left( 1 + \frac{\mathcal{E}}{(b_g + f\mu^2)^2 P} \right)^{-1}, \quad (16)$$

so the deviation of  $r$  from unity is only controlled by the shot noise of the tracer. Again, if  $\mathcal{E} = 0$ , the dark matter power spectrum disappears completely, so that  $b_g$  and  $f$  can be determined without sampling variance.

## 2. Fisher information

In order to determine more quantitatively how much information on cosmology is buried in the clustering statistics of biased tracers and the dark matter, we have to compute the Fisher information matrix [38], a derivation of which is presented in the following: we start with a multivariate Gaussian likelihood of the data vector  $\boldsymbol{\delta}_g$

$$\mathcal{L} = \frac{1}{(2\pi)^{N/2} \sqrt{\det \mathbf{C}}} \exp \left( -\frac{1}{2} \boldsymbol{\delta}_g^\dagger \mathbf{C}^{-1} \boldsymbol{\delta}_g \right), \quad (17)$$

which is a reasonable assumption on large scales, where  $\delta \ll 1$ . The Fisher information matrix for the parameters  $\theta_i$  and  $\theta_j$  is obtained by ensemble averaging over the Hessian of the log-likelihood [39, 40],

$$F_{ij} \equiv - \left\langle \frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j} \right\rangle = \frac{1}{2} \text{Tr} \left( \frac{\partial \mathbf{C}}{\partial \theta_i} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_j} \mathbf{C}^{-1} \right). \quad (18)$$

According to the model from Eq. (7), the derivative of the halo covariance matrix with respect to the parameters is

$$\frac{\partial \mathbf{C}}{\partial \theta_i} = \left( \mathbf{b} \mathbf{b}_i^\dagger + \mathbf{b}_i \mathbf{b}^\dagger \right) P + \mathbf{b} \mathbf{b}_i^\dagger P_i, \quad (19)$$

where  $\mathbf{b}_i \equiv \partial \mathbf{b} / \partial \theta_i$ ,  $P_i \equiv \partial P / \partial \theta_i$  and we assume  $\partial \mathcal{E} / \partial \theta_i = 0$ . Utilizing the *Sherman-Morrison* formula [41, 42], the inverse of the covariance matrix becomes

$$\mathbf{C}^{-1} = \mathcal{E}^{-1} - \frac{\mathcal{E}^{-1} \mathbf{b} \mathbf{b}^\dagger \mathcal{E}^{-1} P}{1 + \mathbf{b}^\dagger \mathcal{E}^{-1} \mathbf{b} P}, \quad (20)$$

provided the shot noise matrix is not singular (e.g., with vanishing  $\mathcal{E}$  also  $\mathbf{C}$  becomes singular by construction). The full Fisher matrix for this case is calculated in Appendix A. The result is

$$F_{ij} = \left[ \Sigma^{-1} \left( \Sigma_{ij} + \frac{\Sigma_i \Sigma_j}{\Sigma} \right) + \left( \Sigma_{ij} - \frac{\Sigma_i \Sigma_j}{\Sigma} \right) + \frac{\Sigma_i}{\Sigma} \frac{P_j}{P} + \frac{\Sigma_j}{\Sigma} \frac{P_i}{P} + \frac{P_i P_j}{2P^2} \right] (1 + \Sigma^{-1})^{-2}, \quad (21)$$

where

$$\Sigma_{ij} \equiv \mathbf{b}_i^\dagger \mathcal{E}^{-1} \mathbf{b}_j P, \quad (22)$$

and  $\Sigma_i$  as well as  $\Sigma$  are defined accordingly by simply omitting the corresponding indices (derivatives). We can identify the first two terms in the square brackets as a *single-tracer* and a *multitracer* term, respectively. The single-tracer term is suppressed by a factor  $\Sigma^{-1}$  as compared to the multitracer term. By definition, the latter vanishes for the case of only one tracer, since then  $\Sigma \Sigma_{ij} = \Sigma_i \Sigma_j$  and Eq. (21) simplifies to

$$F_{ij} = \left( 2 \frac{b_i b_j}{b^2} + \frac{b_i}{b} \frac{P_j}{P} + \frac{b_j}{b} \frac{P_i}{P} + \frac{P_i P_j}{2P^2} \right) \left( 1 + \frac{\mathcal{E}}{b^2 P} \right)^{-2}. \quad (23)$$

While the Fisher information for multiple tracers can in principle become infinite in the limit of no shot noise, Eq. (23) reaches a finite limit when  $\mathcal{E} \rightarrow 0$ .

We consider two sets of parameters separately: the ones that influence only the effective bias,

$$\boldsymbol{\theta}^{(b)} \equiv (\mathbf{b}_g, f) \quad (24)$$

and the ones that go into the matter power spectrum,

$$\boldsymbol{\theta}^{(P)} \equiv (\sigma_8, n_s, h, \Omega_\Lambda, \Omega_m, \Omega_b, \Omega_k). \quad (25)$$

The elements of the shot noise matrix are usually not considered as quantities of interest, but as nuisance parameters that can be marginalized over. From Eq. (21) it is evident that there are degeneracies between  $\boldsymbol{\theta}^{(b)}$  and  $\boldsymbol{\theta}^{(P)}$  due to mixed terms of the form  $\Sigma_i P_j$ . However,

if  $\Sigma$  is sufficiently large, those terms are suppressed in comparison to the term  $\Sigma_{ij}$ .

In case the dark matter density field is known in addition to the galaxies, the derivative of  $\mathbf{C}$  from Eq. (15) becomes

$$\frac{\partial \mathbf{C}}{\partial \theta_i} = \begin{pmatrix} P_i & \mathbf{b}_i^\dagger P + \mathbf{b}^\dagger P_i \\ \mathbf{b}_i P + \mathbf{b} P_i & \mathbf{b} \mathbf{b}_i^\dagger P + \mathbf{b}_i \mathbf{b}^\dagger P + \mathbf{b} \mathbf{b}^\dagger P_i \end{pmatrix}. \quad (26)$$

Furthermore, a block inversion of  $\mathbf{C}$  yields

$$\mathbf{C}^{-1} = \begin{pmatrix} (1 + \Sigma) P^{-1} & -\mathbf{b}^\dagger \mathcal{E}^{-1} \\ -\mathcal{E}^{-1} \mathbf{b} & \mathcal{E}^{-1} \end{pmatrix}, \quad (27)$$

and the Fisher information becomes (see Appendix A)

$$F_{ij} = \Sigma_{ij} + \frac{P_i P_j}{2P^2}. \quad (28)$$

For a single tracer this further simplifies to

$$F_{ij} = \frac{b_i b_j P}{\mathcal{E}} + \frac{P_i P_j}{2P^2}. \quad (29)$$

This expression may increase indefinitely for sufficiently small  $\mathcal{E}$ . In the limit  $\mathcal{E} \rightarrow 0$ , the effective bias can be determined exactly, allowing an exact measurement of the parameters  $\boldsymbol{\theta}^{(b)}$  [12, 43]. On the other hand, constraints on the parameters  $\boldsymbol{\theta}^{(P)}$  are always limited by the variance of the power spectrum  $\text{Var}(P) = 2P^2$ . Here, there are no mixed terms depending both on the effective bias and the power spectrum, so parameter degeneracies between  $\boldsymbol{\theta}^{(b)}$  and  $\boldsymbol{\theta}^{(P)}$  are absent. We note that in the limit of small  $\Sigma_i / \Sigma$ , Eq. (21) reduces to Eq. (28).

### 3. Reparametrization

As mentioned above, the growth rate  $f$  cannot be determined from a galaxy redshift survey alone, since it is degenerate with the power spectrum  $P$ , respectively its normalization  $\sigma_8$ . This degeneracy can only be broken with knowledge of the dark matter density field, or other prior constraints. For this reason it is sometimes convenient to reparametrize by the mapping

$$\tilde{\boldsymbol{\theta}}^{(b)} = \boldsymbol{\theta}^{(b)} / b_g^{\text{ref}} = (\mathbf{b}_g, f) / b_g^{\text{ref}} \equiv (\boldsymbol{\alpha}, \beta), \quad (30)$$

$$\tilde{\sigma}_8 = \sigma_8 b_g^{\text{ref}}, \quad (31)$$

where  $b_g^{\text{ref}}$  is an arbitrary reference galaxy bias so that  $\boldsymbol{\alpha}$  is the relative galaxy bias of all considered tracers to this reference. This mapping leaves the covariance matrix from Eq. (7) unchanged. We can conveniently choose the lowest bias of all tracers as the reference,  $b_g^{\text{ref}} = b_{g1}$ , such that  $\boldsymbol{\alpha} = (1, b_{g2}/b_{g1}, \dots, b_{gN}/b_{g1})$  is always larger than unity and  $\beta = f/b_{g1}$ . Another popular parametrization is

$$\tilde{\boldsymbol{\theta}}^{(b)} = \boldsymbol{\theta}^{(b)} \sigma_8^{\text{ref}} = (\mathbf{b}_g \sigma_8^{\text{ref}}, f \sigma_8^{\text{ref}}), \quad (32)$$

$$\tilde{\sigma}_8 = \sigma_8 / \sigma_8^{\text{ref}}, \quad (33)$$

with  $\sigma_8^{\text{ref}}$  being some arbitrary reference normalization of the power spectrum, which we choose to be identical with our simulation input value of  $\sigma_8 = 0.81$ . In this paper we will quote constraints on both  $\beta$  and  $f\sigma_8$  when considering a galaxy redshift survey only, and on  $f$  when the latter is combined with dark matter observations.

### III. ANALYSIS

#### A. Numerical setup

Our numerical analysis is based on a high-resolution  $N$ -body simulation performed at the University of Zürich supercomputer zBox3 with the GADGET-2 code [44]. It contains  $1536^3$  particles of mass  $4.7 \times 10^{10} h^{-1} M_\odot$  in a box of  $1.3 h^{-1} \text{Gpc}$  a side. We chose our fiducial cosmology to match the WMAP5 best fit with  $\sigma_8 = 0.81$ ,  $n_s = 0.96$ ,  $h = 0.7$ ,  $\Omega_\Lambda = 0.721$ ,  $\Omega_m = 0.279$ ,  $\Omega_b = 0.046$ ,  $\Omega_k = 0$  [45]. We further employ a friends-of-friends algorithm [46] with a linking length of 20% of the mean interparticle distance to generate halo catalogs at different redshifts. With a minimum of 20 dark matter particles per halo we resolve halo masses down to  $M_{\text{min}} \simeq 9.4 \times 10^{11} h^{-1} M_\odot$ , resulting in a mean halo number density of  $\bar{n} \simeq 4.0 \times 10^{-3} h^3 \text{Mpc}^{-3}$  at  $z = 0$  and  $\bar{n} \simeq 3.3 \times 10^{-3} h^3 \text{Mpc}^{-3}$  at  $z = 1$ .

In order to transform the real-space halo catalog into redshift space, we apply Eq. (1) using the velocities of the halos along the three independent directions of the box ( $x$ ,  $y$  and  $z$  axis). Thus, three independent redshift-space catalogs can be constructed from a single real-space catalog, yielding a total effective volume of  $V_{\text{eff}} \simeq 6.6 h^{-3} \text{Gpc}^3$ . Density fields are created by cloud-in-cell interpolation [47] onto a mesh of  $1024^3$  grid points, and the Fourier modes are obtained using a FFT algorithm.

We utilize the IDL algorithm MPFIT [48] to fit our models to the numerical data and find the best-fit parameters including their uncertainties. It is based on the MINPACK distribution by [49] and uses the Levenberg-Marquardt technique to find the minimum of a multidimensional nonlinear least-squares problem. Parameter uncertainties are calculated via the Jacobian of the chi square, which is determined numerically using finite-difference derivatives.

#### B. Optimal tracer selection

In order to fully exploit the benefits of the multitracers approach, the question remains on how to ideally construct different tracers from a given galaxy catalog. In this section we will derive the answer to that question and test it on the basis of  $N$ -body simulations.

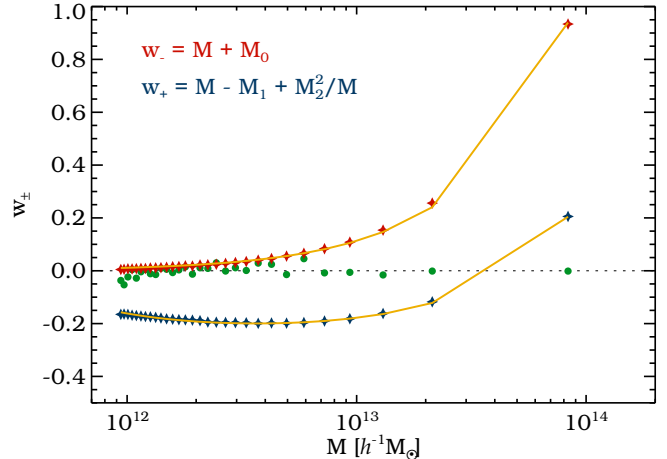


FIG. 1. Three different eigenvectors of the shot noise matrix obtained from a halo catalog that has been split into 30 mass bins of equal number density. The two non-Poisson eigenvectors (upper red and lower blue stars) are overplotted with their best-fit weighting functions  $w_\pm$  as solid lines with functional form given in the top left of the panel in corresponding colors. One representative Poisson eigenvector is shown as green circles; when used as a weighting function it yields a very low clustering signal-to-noise ratio  $\Sigma$ , due to its oscillatory behavior.

##### 1. Principal components of the signal-to-noise ratio

A quantity that plays a crucial role for the cosmological information content contained in the two-point clustering statistics of LSS is the clustering signal-to-noise ratio  $\Sigma \equiv \mathbf{b}^\dagger \mathcal{E}^{-1} \mathbf{b} P$ . For dark matter halos, it has been demonstrated on the basis of numerical  $N$ -body simulations that  $\Sigma$  is dominated by only two *principal components* corresponding to the two nontrivial (non-Poisson) eigenvectors of the shot noise matrix  $\mathcal{E}$  (see Fig. 6 in [50]). We denote these components with a plus and a minus subscript, according to their super- and sub-Poissonian shot noise levels, respectively. In this manner we can expand the clustering signal-to-noise ratio as

$$\Sigma \equiv \mathbf{b}^\dagger \mathcal{E}^{-1} \mathbf{b} P = \sum_{i=0}^N \frac{(\mathbf{V}_i^\dagger \mathbf{b})^2}{\lambda_i} P \simeq \frac{b_-^2}{\mathcal{E}_-} P + \frac{b_+^2}{\mathcal{E}_+} P, \quad (34)$$

where  $\mathbf{V}_i^\dagger$  and  $\lambda_i$  are the  $N$  eigenvectors and eigenvalues of the shot noise matrix  $\mathcal{E}$ , and  $b_\pm$  and  $\mathcal{E}_\pm$  are the effective bias and shot noise of the two principal components of the tracer density field,

$$\delta_\pm \equiv \frac{\sum_i w_\pm(M_i) \delta_g(M_i)}{\sum_i w_\pm(M_i)}, \quad (35)$$

where the summation runs over all individual objects in the volume. The  $w_\pm(M)$  are weighting functions corresponding to the two non-Poisson eigenvectors of the shot

noise matrix, as depicted in Fig. 1. Equation (34) states that a splitting of the tracer density field into  $N$  mass bins yields, in the limit  $N \rightarrow \infty$ , about the same clustering signal-to-noise ratio as simply considering the two weighted fields  $\delta_+$  and  $\delta_-$ . Therefore, these are the most promising candidates to carry out a multitracer analysis with.

In [50] the functional form of  $w_-(M)$  was found to be well described by

$$w_-(M) = M + M_0, \quad (36)$$

where  $M_0$  is about three times the minimum halo mass resolved in the simulation,  $M_0 \simeq 3M_{\min}$ . For the second weighting function we find

$$w_+(M) = M - M_1 + M_2^2/M \quad (37)$$

to be a good fit to the second eigenvector of the shot noise matrix, as shown in Fig. 1. For the constants  $M_1$  and  $M_2$  we find  $M_1 \simeq 1 \times 10^{14} h^{-1} M_\odot$ ,  $M_2 \simeq 3 \times 10^{13} h^{-1} M_\odot$  at  $z = 0$  and  $M_1 \simeq 1 \times 10^{14} h^{-1} M_\odot$ ,  $M_2 \simeq 1 \times 10^{15} h^{-1} M_\odot$  at  $z = 1$ .

The physical origin of the first principal component is related to halo exclusion effects [51]. The sampling of the density field with halos is less stochastic than Poisson sampling (sampling with points) due their finite extension. Since the exclusion volume of halos is proportional to their mass, this effect is strongest at the high mass end [50]. The second principal component can be interpreted as a loop correction to the galaxy bias [52], coming from the second-order term in a local bias expansion model [53]. Due to its nonlinear character, it adds a super-Poissonian shot noise contribution to the two-point clustering statistics of halos, originating from the squared density field [54, 55]. However, through mode coupling it also yields a second-order clustering signal that originates from the bispectrum (three-point function) and adds valuable information coming from smaller scales.

Figure 2 displays the three covariance matrix elements  $C_{++} = \langle \delta_+ \delta_+^* \rangle$ ,  $C_{+-} = \langle \delta_+ \delta_-^* \rangle$  and  $C_{--} = \langle \delta_- \delta_-^* \rangle$  in real space ( $\mu = 0$ ) at redshift  $z = 0$ , extracted from our simulation. The three power spectra are obviously highly correlated and closely follow the shape of the estimated dark matter power spectrum up to  $k \simeq 0.1 h \text{Mpc}^{-1}$ . By construction, the shot noises of the two fields are not correlated, i.e.,  $\langle \epsilon_- \epsilon_+ \rangle = 0$ . Taking the ratios as in Eq. (14) yields three possible estimators for the relative galaxy bias  $\alpha \equiv b_{g-}/b_{g+} \simeq 1.6$ . In the following two subsections, we will provide evidence for the claim that the fields  $\delta_+$  and  $\delta_-$  are indeed the optimal choice for a multitracer analysis.

## 2. Sampling variance cancellation

The idea of utilizing multiple tracers is to cancel sampling variance from the underlying density field  $\delta$ . To

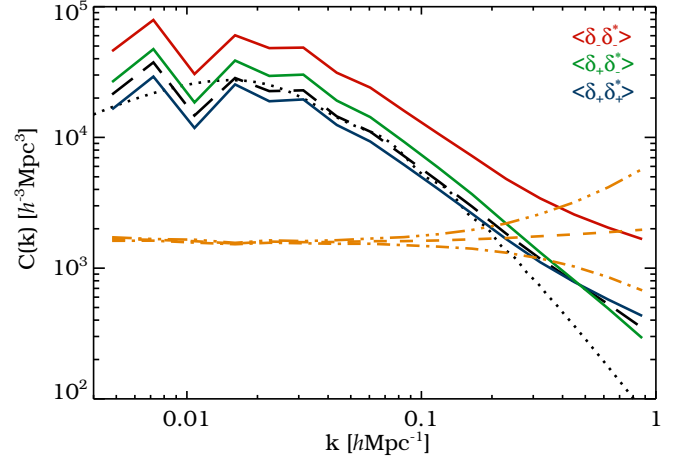


FIG. 2. Real-space auto-power and cross-power spectra of the two fields  $\delta_+$  and  $\delta_-$  (as indicated), obtained through weighting the halo catalog by the two non-Poisson eigenvectors of the shot noise matrix  $w_+$  and  $w_-$ . The long-dashed black line shows the dark matter power spectrum from the simulation and the dotted line its linear theory prediction. The three different estimators for the relative bias  $\alpha$  between  $\delta_+$  and  $\delta_-$ , as defined in Eq. (14), are depicted in dot-dashed ( $C_{+-}/C_{++}$ ), dot-dot-dashed ( $C_{--}/C_{+-}$ ) and dashed ( $\sqrt{C_{--}/C_{++}}$ ). For visibility, they were shifted upwards by a factor of  $10^3$ .

quantify the magnitude of cancellation between any two tracers  $\delta_{g1}$  and  $\delta_{g2}$ , we define the following statistic:

$$\sigma_{\text{SV}}^2 \equiv \frac{\langle |b_2 \delta_{g1}(\mathbf{k}, \mu) - b_1 \delta_{g2}(\mathbf{k}, \mu)|^2 \rangle}{\langle |b_2 \delta_{g1}(\mathbf{k}, \mu)|^2 \rangle + \langle |b_1 \delta_{g2}(\mathbf{k}, \mu)|^2 \rangle}. \quad (38)$$

Here,  $b_1$  and  $b_2$  is the effective bias as defined in Eq. (6). If the two tracers are completely uncorrelated ( $r = 0$ ), there is no cancellation and  $\sigma_{\text{SV}} = 1$ . Yet, if they are perfectly correlated ( $r = 1$ ),  $\sigma_{\text{SV}} = 0$ . Because the real and imaginary parts of any given Fourier mode are uncorrelated, we can swap them for one of the tracers to mimic the case of no correlation.

We consider the following selection criteria for two tracers from our halo catalog:

20/80: lightest 20% vs heaviest 80% of all halos,

50/50: lightest 50% vs heaviest 50% of all halos,

80/20: lightest 80% vs heaviest 20% of all halos,

$u/w_-$ : all uniformly weighted vs all  $w_-$ -weighted halos,

$w_+/w_-$ : all  $w_+$ -weighted vs all  $w_-$ -weighted halos.

The first three are simply obtained via cutting the halo catalog in two at different mass thresholds  $M_{\text{cut}}$  to yield the indicated abundances in each bin. The fourth selection utilizes the whole uniform halo catalog (not weighted) and its  $w_-$ -weighted form, while the last one

TABLE I. Details of the tracer selection.

Selection	$M_{\text{cut}}$ [ $h^{-1}\text{M}_\odot$ ]	$\bar{M}_1$ [ $h^{-1}\text{M}_\odot$ ]	$\bar{M}_2$ [ $h^{-1}\text{M}_\odot$ ]	$\bar{n}_1$ [ $h^3\text{Mpc}^{-3}$ ]	$\bar{n}_2$ [ $h^3\text{Mpc}^{-3}$ ]	$\mathcal{E}_{11}$ [ $h^{-3}\text{Mpc}^3$ ]	$\mathcal{E}_{22}$ [ $h^{-3}\text{Mpc}^3$ ]	$\mathcal{E}_{12}$ [ $h^{-3}\text{Mpc}^3$ ]	$b_{g1}$	$b_{g2}$	$\Sigma_{\text{max}}$
20/80	$1.20 \times 10^{12}$	$1.04 \times 10^{12}$	$7.72 \times 10^{12}$	$8.50 \times 10^{-4}$	$3.15 \times 10^{-3}$	1432	964	448	0.899	0.988	74
50/50	$2.00 \times 10^{12}$	$1.33 \times 10^{12}$	$1.13 \times 10^{13}$	$2.00 \times 10^{-3}$	$2.00 \times 10^{-3}$	911	1243	547	0.896	1.042	73
80/20	$5.00 \times 10^{12}$	$1.97 \times 10^{12}$	$2.22 \times 10^{13}$	$3.15 \times 10^{-3}$	$8.50 \times 10^{-4}$	854	1770	606	0.901	1.217	73
$u/w_-$	-	$6.30 \times 10^{12}$	$6.76 \times 10^{13}$	$4.00 \times 10^{-3}$	$4.00 \times 10^{-3}$	812	51	53	0.969	1.457	2067
$w_+/w_-$	-	$2.29 \times 10^{12}$	$6.76 \times 10^{13}$	$4.00 \times 10^{-3}$	$4.00 \times 10^{-3}$	804	51	7	0.902	1.457	2098

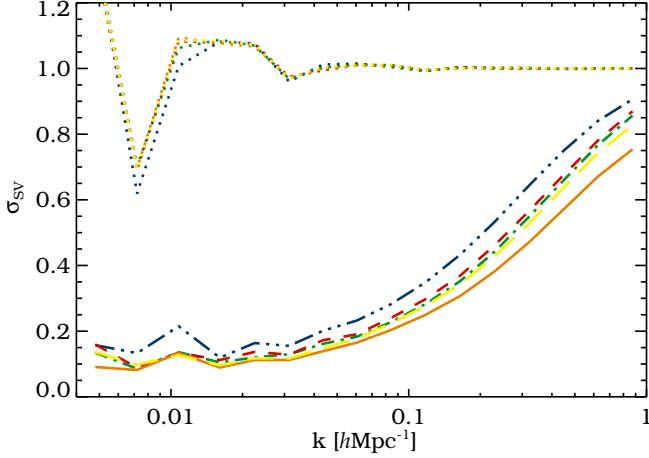


FIG. 3. Sampling variance statistic  $\sigma_{\text{SV}}$  for five different tracer selections at  $z = 0$  as described in the text, 20/80 (dot-dot-dot-dashed, blue), 50/50 (dot-dashed, green), 80/20 (dashed, red),  $u/w_-$  (solid, orange) and  $w_+/w_-$  (long-dashed, yellow). Dotted lines show the corresponding results for uncorrelated modes by swapping the real and imaginary parts of one of the tracer's Fourier modes.

uses both weighting functions  $w_{\pm}$  to construct two tracers from one and the same halo catalog. Further details about the tracers with these selection criteria can be found in Table I.

Figure 3 shows  $\sigma_{\text{SV}}$  for the different tracer pairs; on large scales they all exhibit significant cancellation of sampling variance as compared to the reference case with switched real and imaginary parts (dotted lines). Towards smaller scales this effect is deteriorated due to the onset of nonlinearities and velocity dispersion (Finger-of-God effects) [26]. As evident from the plot, a combination of the uniform with the  $w_-$ -weighted halos yields the highest cancellation of sampling variance. However, the combination of  $w_+$ -weighted and  $w_-$ -weighted halos shows a comparable suppression, as opposed to cutting the halo catalog in two, which yields less cancellation, especially for a low mass cut. If a mass cut is imposed to construct two tracers, the highest sampling variance cancellation is achieved when the same abundance of objects in each of the resulting catalogs is chosen (50/50), which is in agreement with the findings of [56].

### 3. Fisher information

Cancellation of sampling variance alone is not a sufficient indicator on how well cosmological parameters can be constrained in a multitruer analysis. This is because one is looking for relative changes in the clustering signal from multiple tracers, and not for the absolute clustering amplitude in each tracer. In this paragraph we will show that it is desirable to have tracers with a high relative galaxy bias ratio  $\alpha$ .

Quantitatively, the achievable accuracy on a given cosmological parameter  $\theta$  is determined by the inversion of the Fisher matrix of Eq. (21). For the sake of simplicity, let us consider only the  $f$ - $f$  element,

$$F_{ff} = \frac{\Sigma \Sigma_{ff} + \Sigma_f^2 + \Sigma (\Sigma \Sigma_{ff} - \Sigma_f^2)}{1 + \Sigma^2}. \quad (39)$$

In the high signal-to-noise regime,  $\Sigma \gg 1$ , and expressing  $\Sigma$ ,  $\Sigma_f$ , as well as  $\Sigma_{ff}$  in terms of the two principal components as in Eq. (34), we get

$$F_{ff} \simeq \frac{2 \left( \frac{b_-^2}{\mathcal{E}_-^2} + \frac{b_+^2}{\mathcal{E}_+^2} \right) + \frac{(b_- + b_+)^2}{\mathcal{E}_- \mathcal{E}_+}}{\left( \frac{b_-^2}{\mathcal{E}_-^2} + \frac{b_+^2}{\mathcal{E}_+^2} \right)^2} \mu^4 + \frac{\frac{(b_- - b_+)^2}{\mathcal{E}_- \mathcal{E}_+}}{\frac{b_-^2}{\mathcal{E}_-^2} + \frac{b_+^2}{\mathcal{E}_+^2}} \mu^4 P. \quad (40)$$

As a second approximation, we can make use of the fact that  $b_-^2/\mathcal{E}_- \gg b_+^2/\mathcal{E}_+$  and  $\mathcal{E}_- \ll \mathcal{E}_+$  (see Table I), which yields

$$\begin{aligned} F_{ff} &\simeq \frac{2\mu^4}{b_-^2} + \left( 1 - \frac{b_+}{b_-} \right)^2 \frac{\mu^4 P}{\mathcal{E}_+} = \\ &= \frac{2\mu^4}{(b_{g-} + f\mu^2)^2} + \left( 1 - \frac{1 + \beta\mu^2}{\alpha + \beta\mu^2} \right)^2 \frac{\mu^4 P}{\mathcal{E}_+}. \end{aligned} \quad (41)$$

The single-tracer term (as derived in [57]) is dominated by the first principal component of  $\Sigma$ , while the multitruer term depends on the bias ratio  $\alpha = b_{g-}/b_{g+}$  of both principal components and the shot noise  $\mathcal{E}_+$  of the second principle component. In order to maximize  $F_{ff}$ , it is thus desirable to have a large  $\alpha$  and a low  $\mathcal{E}_+$  at the same time. In the special case of uniform Poisson shot noise, Eq. (39) reproduces the expression derived in [58].



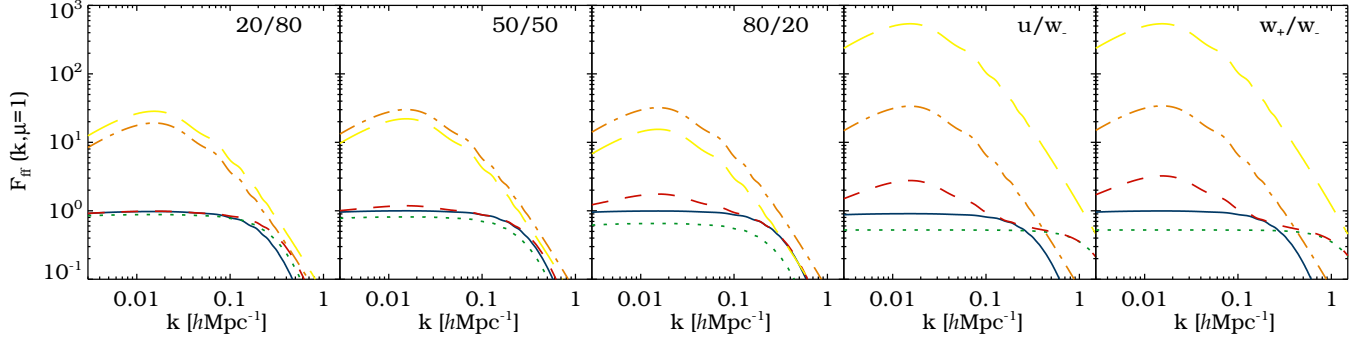


FIG. 4. Fisher matrix element  $F_{ff}$  as a function of  $k$  at  $\mu = 1$  and  $z = 0$  for various tracer selections (indicated in the top right of each panel). Results are shown for a single-tracer analysis with the first (solid, blue) and the second tracer (dotted, green) used separately, both of the tracers combined in a multitracer analysis (dashed, red), and each of the tracers combined with idealistic dark matter observations (dot-dashed orange and long-dashed yellow, respectively).

If the dark matter density field is known separately, this Fisher matrix element becomes

$$F_{ff} = \Sigma_{ff} \simeq \frac{\mu^4 P}{\mathcal{E}_-}, \quad (42)$$

thus independent of the effective bias and limited only by the low shot noise level of the first principal component.

Figure 4 depicts  $F_{ff}(k, \mu = 1)$  for all of our five tracer selections. For each case we further distinguish between the following scenarios:

- single-tracer analysis with each of the two tracers taken separately,
- multiple-tracer analysis with both tracers combined,
- combined analysis of each tracer with the dark matter density field.

Clearly, in a single-tracer analysis the tracer with the lowest galaxy bias yields the highest information on the growth rate  $f$ . This is evident from Eq. (23), where  $b_g$  enters in the denominator of the first term and the shot noise  $\mathcal{E}$  is negligible if  $\mathcal{E}/b_g^2 P \ll 1$ . Since  $b_{g1}$  is very similar in all five cases, the Fisher information from a single tracer cannot be increased much by any particular choice of tracer.

On the contrary, a combination of two tracers can cancel out sampling variance and therefore considerably increase the available information on  $f$  if the tracers are selected appropriately. As evident from the dashed red lines in Fig. 4, the multitracer term in Eq. (41) gains importance over the single-tracer term as moving from the left to the right panel. Again, the two tracers obtained by a low mass cut yield the worst results, owing to the fact that the shot noise of both tracers is super-Poissonian in this selection. It is more optimal to impose a high mass cut in order to benefit from the sub-Poissonian shot noise level of the heaviest halos [50]. However, the highest

Fisher information content is obtained when correlating the two orthogonally weighted fields  $\delta_+$  and  $\delta_-$ . The main reason for this is the large relative galaxy bias between the two fields and their relatively low shot noise level [see Table I and Eq. (41)].

We have also explored the possibility of splitting the halo catalog into more than two mass bins, considering up to  $N = 10$  tracers. We find that the Fisher information from multiple tracers increases with the number of bins  $N$  and approaches the result obtained with the two weighted fields  $\delta_+$  and  $\delta_-$  in the limit of high  $N$ .

Finally, adding the information from the dark matter density field to each one of the tracers increases the information content on  $f$ . In this case, sampling variance inherent in the density field  $\delta$  is known and can thus be removed from the halo fields directly. According to Eq. (42), the Fisher information is inversely proportional to the shot noise of the tracer, so the lowest stochasticity weight  $w_-$  yields the best results.

### C. Multitracer fit

So far we have investigated the Fisher information content on the growth rate using multiple tracers of the LSS. The question of how to actually constrain parameters of interest from a data set in the optimal way will be answered in this section. For this task we want to maximize the likelihood function from Eq. (17), which is equivalent to minimizing its negative logarithm, the chi square

$$\chi^2 \equiv \sum_{\mathbf{k}} \frac{1}{2} \delta_g^\dagger(\mathbf{k}, \mu) \mathbf{C}^{-1} \delta_g(\mathbf{k}, \mu) + \frac{1}{2} \ln(\det \mathbf{C}) . \quad (43)$$

Here, the covariance matrix  $\mathbf{C}$  is given by the clustering model of Eq. (7) and we sum over all individual Fourier modes  $\delta_g(\mathbf{k}, \mu)$  from our halo catalog. When we add the dark matter density field as an observable, we use the model from Eq. (15) and  $\delta = (\delta, \delta_g)$  as our data vector.

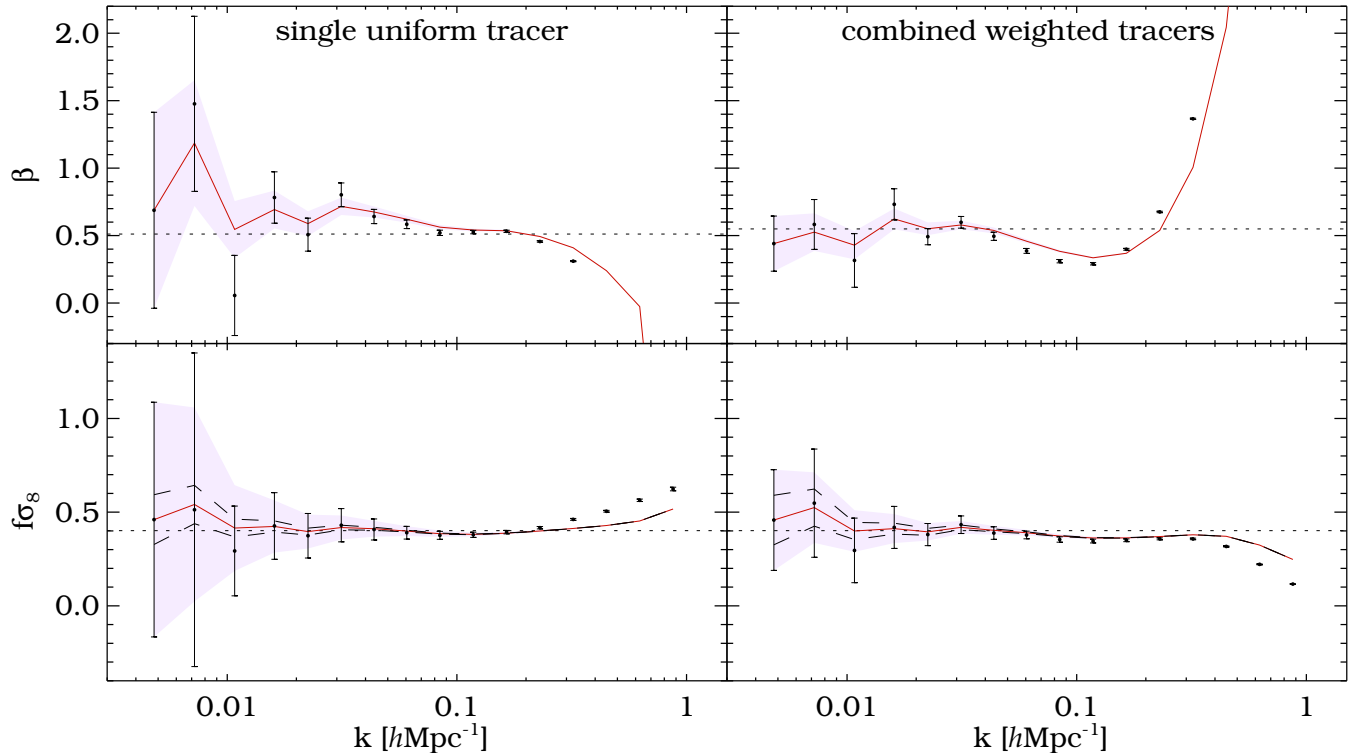


FIG. 5. Fit for the redshift-space distortion parameter  $\beta$  (top) and the product  $f\sigma_8$  (bottom) from the two-point clustering statistics of halos in an  $N$ -body simulation with effective volume  $V_{\text{eff}} \simeq 6.6h^{-3}\text{Gpc}^3$  and halo-mass resolution  $M_{\text{min}} \simeq 9.4 \times 10^{11}h^{-1}\text{M}_{\odot}$  at  $z = 0$ . LEFT: Conventional single-tracer analysis utilizing all halos (not weighted) from the same catalog. RIGHT: Multitracer analysis with the two fields  $\delta_+$  and  $\delta_-$ , obtained through weighting the halo catalog with its principal components  $w_+$  and  $w_-$ . The fitting results are shown in logarithmic bins of  $k$  (points with  $1\sigma$  error bars), as well as cumulative as a function of  $k = k_{\text{max}}$  with fixed  $k_{\text{min}} = 0.0048h\text{Mpc}^{-1}$  (red solid lines with shaded region). Dotted lines show the linear theory prediction with  $f = \Omega_m^{0.55}$  and  $\sigma_8 = 0.81$ , dashed lines the cumulative sampling variance limit.

Figure 5 presents the fitting results for the RSD parameter  $\beta$  and the product of growth rate  $f$  with the normalization of the power spectrum  $\sigma_8$  from our halo catalogs at redshift  $z = 0$ . While  $\beta$  has been obtained from a single-parameter fit, we have marginalized over the galaxy bias of each tracer as a free parameter in the fit for  $f\sigma_8$ . In the left column, the standard single-tracer analysis utilizing all objects in the halo catalog is performed. The best fits along with their  $1\sigma$ -error bars are shown both in  $k$  bins (points with error bars), as well as cumulative as a function of  $k = k_{\text{max}}$  with fixed  $k_{\text{min}} = 0.0048h\text{Mpc}^{-1}$  (solid lines with shaded region).

The constraints on  $\beta$  and  $f\sigma_8$  are clearly affected by sampling variance, as evident from the large scatter of the points at low  $k$ . When the number of available Fourier modes grows towards higher  $k$ , this scatter becomes smaller; however, beyond a scale of  $k \simeq 0.2h\text{Mpc}^{-1}$ , linear theory breaks down and the fits depart from their scale-independent linear value assuming  $f = \Omega_m^{0.55}$  (dotted line). The cumulative sampling variance limit for the determination of  $f\sigma_8$  is shown as a dashed line. Clearly, the single-tracer fit yields a substantially larger uncer-

tainty compared to this limit.

When combining the two  $w_+/w_-$ -weighted tracers in a multitracer analysis as shown in the right column of Fig. 5, the scatter of the fit is significantly suppressed. On the largest scales, the errors are reduced by up to a factor of 4 and the constraints on  $f\sigma_8$  reach the sampling variance limit closely. However, the fit seems to deviate from the linear theory prediction already at  $k \simeq 0.04h\text{Mpc}^{-1}$ . This is likely due to the high bias of the  $w_-$ -weighted tracer: as shown in [59], more highly biased halos show a stronger scale dependence in redshift space, invalidating the Kaiser formula on even larger scales. This could be corrected for by nonlinear RSD models, which is beyond the scope of this paper. On the other hand, the apparently more linear behavior of the single-tracer analysis may likely be coincidental at  $k > 0.04h\text{Mpc}^{-1}$ . This is supported by results shown in the left panel of Fig. 6. Here, we combine the uniform halo catalog (without weighting) with the dark matter field from our simulation to fit for  $f$  directly. Obviously, sampling variance has decreased even further, but deviations from linear theory already kick in at a scale of  $k \simeq 0.04h\text{Mpc}^{-1}$ .

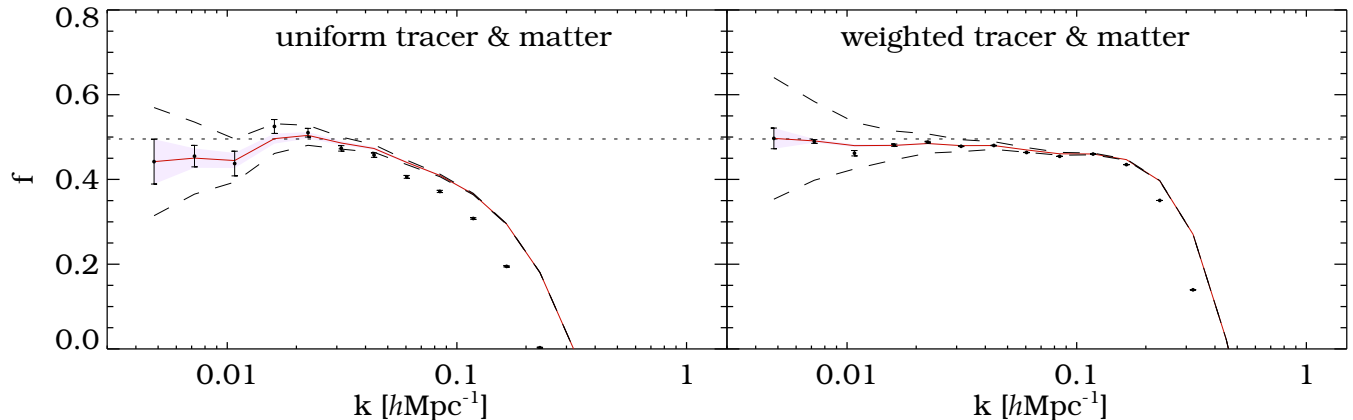


FIG. 6. Fit for the growth rate  $f$  from the two-point clustering statistics of halos *and* the dark matter combined. LEFT: All uniform halos (not weighted) and the dark matter. RIGHT: All halos, weighted with the lowest stochasticity weight  $w_-$ , and the dark matter. The meaning of lines and symbols is the same as in Fig. 5.

Most impressive constraints on the growth rate are obtained when we combine the  $w_-$ -weighted halos with the dark matter density field, as depicted in the right panel of Fig. 6. In this case, sampling variance has almost canceled out completely and the error bars on  $f$  have diminished by up to a factor of 142 (at the peak of the power spectrum), when compared to the standard single-tracer analysis. Moreover, deviations from linear theory are very small up to scales of  $k \sim 0.2 h\text{Mpc}^{-1}$ , making this kind of experiment the most promising one out of the four considered scenarios.

Methods to combine galaxy clustering and weak lensing data have been studied extensively in the recent literature (e.g., [60–66]), suggesting high improvements for precision cosmology. Unfortunately, obtaining the dark matter density field in 3D from observations is a highly nontrivial problem and is subject of active research. Weak lensing tomography is the technique aiming to achieve this goal (e.g., [67–71]), but the resolution in the radial direction is not expected to be high because of the relatively broad lensing kernels along the line of sight. Moreover, we assume an ideal reconstruction of the dark matter density field without considering additional sources of error involved in the lensing measurement, such as *shape noise* and *intrinsic alignment*, for example.

Without knowledge about  $\delta$ , in principle we would have to marginalize over all the parameters  $\theta^{(P)}$  of the dark matter power spectrum as well. This implies calculating the transfer function in each iteration of the fitting procedure, which goes beyond the scope of this paper. In the high signal-to-noise regime of the multitracers analysis, the degeneracy between the parameters  $\theta^{(P)}$  and  $\theta^{(b)}$  is expected to be rather weak (except the fundamental degeneracy between  $f$  and  $\sigma_8$ ), as the mixed terms between  $\Sigma_i$  and  $P_j$  in Eq. (21) are suppressed by  $\Sigma$ . Moreover, in

Appendix C of [52] it has been shown that  $P(k)$  cancels out to a high degree in the chi square of Eq. (43). Of course, in case the dark matter density field is available, we do not have to worry about those issues, since  $b_g$  and  $P$  are directly observable.

In order to quantify the gains in accuracy, we compare the size of the error bars on  $\beta$  and  $f\sigma_8$  from the multitracers to the single-tracer analysis in Fig. 7 (solid and dashed blue lines, respectively). Obviously, sampling variance mostly cancels on the largest scales, yielding improvements in accuracy of up to a factor of 4. Beyond scales of  $k \simeq 0.1 h\text{Mpc}^{-1}$ , the improvement is deteriorated due to the onset of nonlinear clustering and mode coupling. Deviations from linear theory increase towards smaller scales, making the fit of this model to the data increasingly biased (see Figs. 5 and 6). Therefore, the drop of the curves at  $k > 0.3 h\text{Mpc}^{-1}$  is likely an artifact of the fitting procedure using an incorrect model and should not be trusted. At redshift  $z = 1$  this turnover is moved to smaller scales, the overall improvement in the error ratio is, however, slightly deteriorated. In order to access the cosmological information content buried in the semilinear regime of galaxy clustering in redshift space, one cannot avoid having to invoke more elaborate models involving perturbative methods, such as the ones proposed in [72].

More gains can be achieved when galaxies and ideal dark matter observations are combined, the improvement in the accuracy on  $\beta$  compared to the ordinary single-tracer analysis amounts to about a factor of 10 in this case (solid green line). If, additionally, galaxies are weighted optimally, it increases by another factor of 10, two orders of magnitude better than what a single-tracer analysis can achieve. In this case the improvement even persists down to smaller scales of  $k \simeq 0.3 h\text{Mpc}^{-1}$ . However, the effect of optimal weighting is diminished towards higher redshifts.

Unfortunately, the halo masses we used to construct

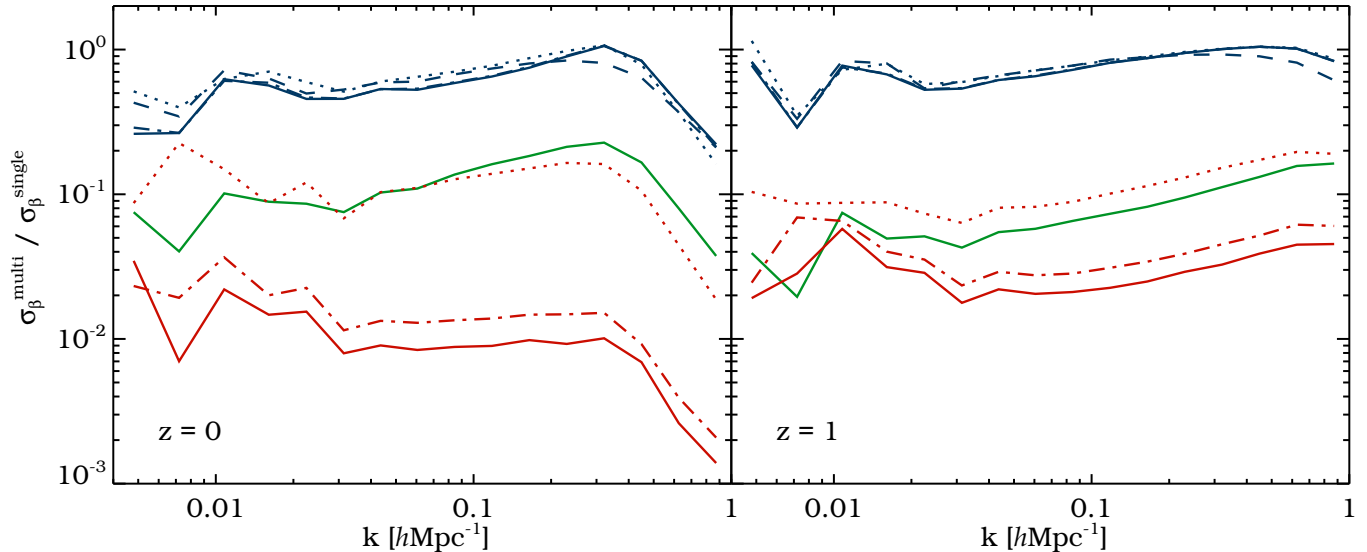


FIG. 7. Ratios of the binned one-sigma error bars (from Fig. 5) on  $\beta$  (solid blue) and  $f\sigma_8$  (dashed blue) when comparing the optimal multitracer to the single-tracer analysis at  $z = 0$  (left panel) and  $z = 1$  (right panel). Additionally, the improvement in constraints on  $\beta$  from a combined clustering analysis of halos *and* dark matter as compared to the single-tracer case is shown in solid green (no weighting of halos) and in solid red (optimally weighted halos). Adding a log-normal scatter of  $\sigma_{\ln M} = 0.1$  (dot-dashed) and  $\sigma_{\ln M} = 0.5$  (dotted) to the halo masses results in a degradation of the constraints from the weighted fields.

our weighted density fields are not directly observable in reality. Yet, they correlate with many observables, such as X-ray luminosity, galaxy richness, weak lensing shear, velocity dispersion or the thermal Sunyaev-Zel'dovich (SZ) effect (e.g., [73]). Scaling relations between these observables and halo mass can be calibrated with numerical simulations to obtain unbiased mass proxies with minimal scatter (e.g., [74, 75]). While optical methods show a rather large scatter of  $\sigma_{\ln M} \simeq 0.45$  [76], X-ray or SZ observations yield tighter relations with  $\sigma_{\ln M}$  reaching below 0.1 [77].

We artificially add a constant log-normal scatter to the halo masses of our catalog in order to mimic the observational uncertainties in the mass determination for a rather pessimistic scenario ( $\sigma_{\ln M} = 0.5$ ) and a more optimistic scenario ( $\sigma_{\ln M} = 0.1$ ). The results are depicted in Fig. 7 as dotted and dot-dashed lines, respectively. While the constraints from the multitracer analysis on  $\beta$  are only marginally affected for  $\sigma_{\ln M} = 0.1$  and degrade by roughly 20–30% for  $\sigma_{\ln M} = 0.5$ , the combined analysis using optimally weighted halos and the dark matter is more severely deteriorated by mass scatter. Here, even the optimistic scenario increases the uncertainty on the growth rate by roughly 50%, while the benefits from optimal weighting are completely lost when going to  $\sigma_{\ln M} = 0.5$ .

Clearly, the high level of precision that can be obtained with this kind of experiment demands precise mass estimates. Fortunately, a whole industry of existing and planned experiments devoted to cluster cosmology will

provide those high quality data (e.g., [78–83]).

#### D. Halo model predictions

In this section we want to investigate how our results depend on the resolution of the simulation; respectively, the minimum resolved halo mass. In a real experiment, this corresponds to the depth of the galaxy survey with a corresponding luminosity threshold. Because  $N$ -body simulations with a given volume become increasingly expensive with higher resolution, we will turn to theoretical predictions henceforth.

The clustering properties of dark matter and halos can be neatly described by the *halo model* (see, e.g., [84]). The basic idea is to separately describe the clustering within a given halo (one-halo term) and the clustering amongst different halos (two-halo term). In [50] the halo model is utilized to derive an analytical expression for the shot noise matrix. The result can be written as

$$\mathcal{E} = \bar{n}^{-1} \mathbf{I} - \mathbf{b}_g \mathcal{M}^\dagger - \mathcal{M} \mathbf{b}_g^\dagger, \quad (44)$$

where  $\bar{n}$  is the number density of halos per bin,  $\mathbf{I}$  the identity matrix,  $\mathcal{M} \equiv \mathbf{M}/\bar{\rho}_m - \mathbf{b}_g \langle n M^2 \rangle / 2 \bar{\rho}_m^2$  and  $\mathbf{M}$  a vector containing the mean halo mass of each bin. In Appendix B we utilize this expression to derive the clustering signal-to-noise ratio  $\Sigma$  (as well as  $\Sigma_i$  and  $\Sigma_{ij}$ ) from the halo model. With Eq. (39) we can then determine the Fisher information on the growth rate and compare the single-tracer analysis to the multitracer analysis.

Figure 8 displays the same ratio as Fig. 7 for the uncertainty on  $\beta$ , but now as a function of minimum halo mass  $M_{\min}$  at a fixed scale of  $k \simeq 0.016 h \text{Mpc}^{-1}$  (peak of the power spectrum). The gains from the multitracers method kick in at  $M_{\min} \simeq 10^{14} h^{-1} \text{M}_{\odot}$ , where  $\Sigma \simeq 1$ , and increase towards lower  $M_{\min}$  due to the growing signal-to-noise ratio.

A combination of galaxy and dark matter observations may increase these gains further, especially when the galaxies are weighted optimally. In this case there is no saturation towards lower  $M_{\min}$  and the improvement compared to the single-tracer analysis continues to grow. In contrast, the uniform galaxy overdensity field (not weighted) combined with dark matter already shows up a saturation at  $M_{\min} \simeq 10^{10} h^{-1} \text{M}_{\odot}$ , so no more information on the growth rate can be gained when even more lighter halos are included in this kind of analysis. Also note that the halo model predictions underestimate the improvements obtained when adding dark matter clustering information, as our  $N$ -body results with  $M_{\min} \simeq 9.4 \times 10^{11} h^{-1} \text{M}_{\odot}$  yield higher improvements (see Fig. 7).

In the shot noise dominated regime above a minimum mass of  $M_{\min} \sim 10^{14} h^{-1} \text{M}_{\odot}$  the error ratio decreases again towards higher  $M_{\min}$  because the Fisher information on  $f$  from a single tracer here roughly scales as  $\mathcal{E}^{-2}$ , while for a tracer combined with dark matter as  $\mathcal{E}^{-1}$  [see Eqs. (23) and (29), respectively]. We refer the reader to Figs. 10 and 11 of [52], where a similar plot is shown for the individual error bars on the non-Gaussianity parameter  $f_{\text{NL}}$ .

#### IV. CONCLUSIONS

In this paper we investigated the benefits of using weights in a multitracers analysis of LSS, with a particular focus on constraining the growth rate of structure formation. On the basis of earlier results on the clustering properties of dark matter halos and their stochasticity [50], we argue that the gains from a multitracers analysis in the sense of [12] can be achieved by considering only the two principal components of the clustering signal-to-noise ratio  $\Sigma$  (or, equivalently, the two non-Poisson eigenvectors of the shot noise matrix  $\mathcal{E}$ ).

We present their explicit functional forms in terms of weights, showing that the first one coincides with the weighting function explored in previous work [50], giving rise to low stochasticity and high bias. For the second one the weights are also mass dependent, but have a zero crossing, such that the overall bias is low. This yields a high relative galaxy bias  $\alpha$  between the two tracers, maximizing the Fisher information content on the cosmological parameters [12]. All of the other eigenvectors oscillate around zero and add very little information. The advantage of reducing the information to two eigenvec-

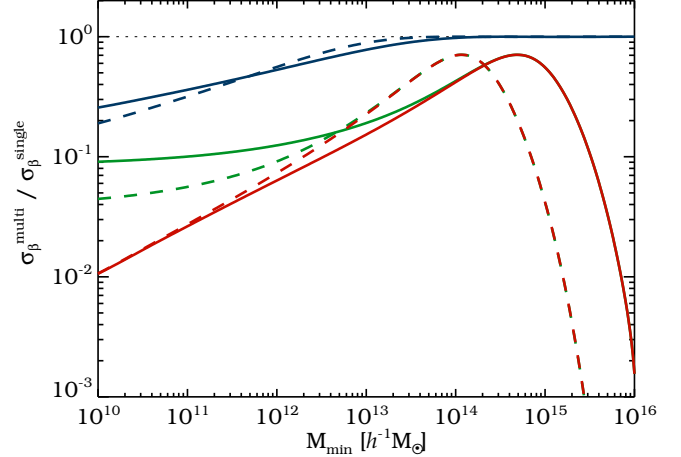


FIG. 8. Halo model prediction for the improvement on  $\sigma_{\beta}$  from the single-tracer analysis to the optimal multitracers analysis (blue) at  $k \simeq 0.016 h \text{Mpc}^{-1}$  as a function of the lower halo-mass threshold  $M_{\min}$  at  $z = 0$  (solid) and  $z = 1$  (dashed). The results for the combined analysis of uniform halos (not weighted) and dark matter (green), as well as the combination of optimally weighted halos with the dark matter (red) are also depicted.

tors is that all of the objects in a given catalog can be used to construct the two principal components, while in the conventional multitracers analysis the catalog has to be split into many lower number density subsamples with higher shot noise [56, 58].

On the basis of numerical  $N$ -body simulations of dark matter halos, we demonstrate that the constraints on  $\beta$  and  $f\sigma_8$  can be improved by up to a factor of 4 relative to a single-tracer method, but most of the improvement comes from large scales (low  $k$ ), while for higher  $k$  the gains are smaller and vanish above  $k \sim 0.1 h \text{Mpc}^{-1}$ , where nonlinear effects introduce additional stochasticity between the two tracers. This technique is fairly insensitive to the observational uncertainty on the halo masses, as even a 50% log-normal scatter does not degrade the improvements significantly. Halo model considerations suggest even higher gains of the method with increasing mass resolution.

One potential concern for our method is the possibility that galaxies might be bad tracers of their host-halo centers [85] and therefore exhibit less pronounced principal components in the clustering signal-to-noise ratio that are distinct from Poisson sampling. However, there are strong indications that certain types of galaxies do show strong correlations in both position and mass with their host halo, e.g. luminous red galaxies (LRGs) [86]. Techniques to distinguish satellite galaxies from central galaxies have been developed and the satellite fraction can be used as an estimator of the host-halo mass [87]. Mock LRG catalogs obtained from a halo occupation distribution (HOD) model suggest some reduction in stochastic-

ity is possible even without explicit knowledge of halo masses [88]. Therefore, an achievement of the presented gains seems feasible in light of upcoming spectroscopic galaxy surveys such as EUCLID [3], which will attain galaxy number densities, host-halo mass ranges and a survey volume comparable to the simulations used in this paper [89, 90].

Whether these gains translate into a useful constraint on the final cosmological parameters depends on our ability to model nonlinear RSD effects. We find nonlinear effects are important for  $\beta$  already at  $k > 0.03h\text{Mpc}^{-1}$ , although they appear to be important for  $f\sigma_8$  only at  $k > 0.1h\text{Mpc}^{-1}$ . In the most pessimistic case where the RSD model cannot be trusted for  $k > 0.03h\text{Mpc}^{-1}$ , the multitracers method provides major gains relative to the single-tracer case, but neither method provides very strong constraints overall because of the limited number of available Fourier modes. In the case where we can use all the modes up to  $k \sim 0.1h\text{Mpc}^{-1}$ , the overall errors are considerably smaller and the multitracers method provides less of an advantage. It is clear that a better modeling of the nonlinear effects in RSD is needed to understand the ultimate reach of RSD in both single-tracer and multitracers methods.

In a more idealistic scenario, we also consider the joint analysis of halos and the dark matter density field, which in principle is achievable via a combination of spectroscopic redshift surveys and weak lensing tomography. Here, utilizing optimal weights can yield up to two orders of magnitude improvements in constraining  $\beta$  as compared to a single-tracer analysis, but the method is more prone to uncertainties in the halo mass estimates. It is unlikely that this gain can be achieved in practice, since it is very difficult to measure dark matter clustering in the radial direction directly.

A further technique to construct differently biased tracers of the density field makes use of nonlinear transformations [91]. Although it is difficult to describe the effects of a nonlinear transformation on both signal and noise in galaxy clustering data, combined with optimal weights this may provide another tool for the multitracers analysis.

In this paper we have focused on the information that can be extracted from RSD, in particular  $\beta$  and  $f\sigma_8$ , but our method is not limited to constraints on the growth rate, but may be applied to the analysis of primordial non-Gaussianity [52], general relativistic corrections in large-scale clustering [92], the Alcock-Paczyński test [12] or any other quantity that influences the effective bias of tracers of the density field. It is possible that a better model of nonlinear RSD may yield a more efficient multitracers method, where the gains relative to the single-tracer analysis described here on large scales can be extended to smaller scales. We leave these directions for the future.

## ACKNOWLEDGMENTS

We thank Jaiyul Yoo, Jonathan Blazek, Tobias Baldauf and Zvonimir Vlah for fruitful discussions and Volker Springel for making public his  $N$ -body code GADGET-2. This work is supported by the Packard Foundation, the Swiss National Foundation under Contract No. 200021-116696/1 and WCU Grant No. R32-10130. V.D. acknowledges support by the Swiss National Science Foundation. N.H. appreciated the hospitality of Lawrence Berkeley National Lab and the Institute for the Early Universe at Ewha University Seoul while parts of this work were completed.

### Appendix A: FISHER MATRIX FOR MULTIPLE BIASED TRACERS

With Eqs. (19) and (20) plugged into Eq. (18), we have

$$\begin{aligned}
F_{ij} &= \frac{1}{2} \text{Tr} \left[ \left( \mathbf{b} \mathbf{b}_i^\dagger + \mathbf{b}_i \mathbf{b}^\dagger + \mathbf{b} \mathbf{b}^\dagger \frac{P_i}{P} \right) \mathbf{C}^{-1} \left( \mathbf{b} \mathbf{b}_j^\dagger + \mathbf{b}_j \mathbf{b}^\dagger + \mathbf{b} \mathbf{b}^\dagger \frac{P_j}{P} \right) \mathbf{C}^{-1} \right] P^2 = \\
&= \frac{1}{2} \left( \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b} \mathbf{b}_i^\dagger \mathbf{C}^{-1} \mathbf{b}_j + \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b} \mathbf{b}_j^\dagger \mathbf{C}^{-1} \mathbf{b}_i + \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b}_i \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b}_j + \mathbf{b}_i^\dagger \mathbf{C}^{-1} \mathbf{b} \mathbf{b}_j^\dagger \mathbf{C}^{-1} \mathbf{b} + \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b} \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b}_i \frac{P_j}{P} + \right. \\
&\quad \left. \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b} \mathbf{b}_i^\dagger \mathbf{C}^{-1} \mathbf{b} \frac{P_j}{P} + \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b} \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b}_j \frac{P_i}{P} + \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b} \mathbf{b}_j^\dagger \mathbf{C}^{-1} \mathbf{b} \frac{P_i}{P} + \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b} \mathbf{b}^\dagger \mathbf{C}^{-1} \mathbf{b} \frac{P_i P_j}{P^2} \right) P^2 = \\
&= \frac{\Sigma}{1 + \Sigma} \left[ \Sigma_{ij} - \frac{\Sigma_i \Sigma_j}{1 + \Sigma} \right] + \frac{\Sigma_i \Sigma_j}{(1 + \Sigma)^2} + \frac{\Sigma_i \Sigma}{(1 + \Sigma)^2} \frac{P_j}{P} + \frac{\Sigma_j \Sigma}{(1 + \Sigma)^2} \frac{P_i}{P} + \left( \frac{\Sigma}{1 + \Sigma} \right)^2 \frac{P_i P_j}{2P^2} = \\
&= \left[ \Sigma^{-1} \left( \Sigma_{ij} + \frac{\Sigma_i \Sigma_j}{\Sigma} \right) + \left( \Sigma_{ij} - \frac{\Sigma_i \Sigma_j}{\Sigma} \right) + \frac{\Sigma_i}{\Sigma} \frac{P_j}{P} + \frac{\Sigma_j}{\Sigma} \frac{P_i}{P} + \frac{P_i P_j}{2P^2} \right] (1 + \Sigma^{-1})^{-2} . \quad (\text{A1})
\end{aligned}$$

With additional knowledge about the dark matter density field we need to work out Eq. (18) by plugging in Eqs. (26) and (27). Let us first note that

$$\frac{\partial \mathbf{C}}{\partial \theta_i} \mathbf{C}^{-1} = \begin{pmatrix} P_i/P - \Sigma_i & \mathbf{b}_i^\dagger \boldsymbol{\mathcal{E}}^{-1} P \\ \mathbf{b}_i + \mathbf{b} (P_i/P - \Sigma_i) & \mathbf{b} \mathbf{b}_i^\dagger \boldsymbol{\mathcal{E}}^{-1} P \end{pmatrix} ,$$

so

$$\begin{aligned}
&\frac{\partial \mathbf{C}}{\partial \theta_i} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_j} \mathbf{C}^{-1} = \\
&\begin{pmatrix} (P_i/P - \Sigma_i)(P_j/P - \Sigma_j) + \Sigma_{ij} + \Sigma_i(P_j/P - \Sigma_j) & (P_i/P - \Sigma_i) \mathbf{b}_j^\dagger \boldsymbol{\mathcal{E}}^{-1} P + \Sigma_i \mathbf{b}_j^\dagger \boldsymbol{\mathcal{E}}^{-1} P \\ [\mathbf{b}_i + \mathbf{b} (P_i/P - \Sigma_i)](P_j/P - \Sigma_j) + \mathbf{b} \Sigma_{ij} + \mathbf{b} \Sigma_i (P_j/P - \Sigma_j) & \mathbf{b}_i \mathbf{b}_j^\dagger \boldsymbol{\mathcal{E}}^{-1} P + \mathbf{b} \mathbf{b}_j^\dagger \boldsymbol{\mathcal{E}}^{-1} P (P_i/P - \Sigma_i) + \mathbf{b} \Sigma_i \mathbf{b}_j^\dagger \boldsymbol{\mathcal{E}}^{-1} P \end{pmatrix}
\end{aligned}$$

This yields

$$F_{ij} = \frac{1}{2} \text{Tr} \left( \frac{\partial \mathbf{C}}{\partial \theta_i} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_j} \mathbf{C}^{-1} \right) = \Sigma_{ij} + \frac{P_i P_j}{2P^2} . \quad (\text{A2})$$

### Appendix B: HALO MODEL PREDICTION FOR THE CLUSTERING SIGNAL-TO-NOISE RATIO

In the halo model the shot noise matrix is given by Eq. (44). In order to invert  $\boldsymbol{\mathcal{E}}$ , we write  $\boldsymbol{\mathcal{E}} = \boldsymbol{\mathcal{A}} - \mathbf{M} \mathbf{b}_g^\dagger$  with  $\boldsymbol{\mathcal{A}} \equiv \bar{n}^{-1} \mathbf{I} - \mathbf{b}_g \mathbf{M}^\dagger$  and apply the Sherman-Morrison formula twice:

$$\boldsymbol{\mathcal{E}}^{-1} = \boldsymbol{\mathcal{A}}^{-1} + \frac{\boldsymbol{\mathcal{A}}^{-1} \mathbf{M} \mathbf{b}_g^\dagger \boldsymbol{\mathcal{A}}^{-1}}{1 - \mathbf{b}_g^\dagger \boldsymbol{\mathcal{A}}^{-1} \mathbf{M}} , \quad \boldsymbol{\mathcal{A}}^{-1} = \bar{n} \mathbf{I} + \frac{\mathbf{b}_g \mathbf{M}^\dagger \bar{n}}{\bar{n}^{-1} - \mathbf{M}^\dagger \mathbf{b}_g} . \quad (\text{B1})$$

With

$$\Sigma_{ij} \equiv \mathbf{b}_i^\dagger \boldsymbol{\mathcal{E}}^{-1} \mathbf{b}_j P = \frac{\mathbf{b}_i^\dagger \boldsymbol{\mathcal{A}}^{-1} \mathbf{b}_j \left( 1 - \mathbf{b}_g^\dagger \boldsymbol{\mathcal{A}}^{-1} \mathbf{M} \right) + \mathbf{b}_i^\dagger \boldsymbol{\mathcal{A}}^{-1} \mathbf{M} \mathbf{b}_g^\dagger \boldsymbol{\mathcal{A}}^{-1} \mathbf{b}_j}{1 - \mathbf{b}_g^\dagger \boldsymbol{\mathcal{A}}^{-1} \mathbf{M}} P , \quad (\text{B2})$$

and

$$\mathbf{b}_i^\dagger \boldsymbol{\mathcal{A}}^{-1} \mathbf{b}_j = \frac{\mathbf{b}_i^\dagger \mathbf{b}_j \left( \bar{n}^{-1} - \mathbf{M}^\dagger \mathbf{b}_g \right) + \mathbf{b}_i^\dagger \mathbf{b}_g \mathbf{M}^\dagger \mathbf{b}_j}{\bar{n}^{-1} - \mathbf{M}^\dagger \mathbf{b}_g} \bar{n} , \quad (\text{B3})$$

$$1 - \mathbf{b}_g^\dagger \boldsymbol{\mathcal{A}}^{-1} \mathbf{M} = \frac{\left( \bar{n}^{-1} - \mathbf{M}^\dagger \mathbf{b}_g \right)^2 - \mathbf{b}_g^\dagger \mathbf{b}_g \mathbf{M}^\dagger \mathbf{M}}{\bar{n}^{-1} - \mathbf{M}^\dagger \mathbf{b}_g} \bar{n} , \quad (\text{B4})$$



and similar terms combining  $\mathbf{b}$ ,  $\mathbf{b}_i$ ,  $\mathbf{b}_g$  and  $\mathcal{M}$ , after some algebra we get

$$\Sigma_{ij} = \mathbf{b}_i^\dagger \mathbf{b}_j \bar{n} P + \frac{\mathbf{b}_i^\dagger \mathcal{M} \mathbf{b}_g^\dagger \mathbf{b}_g \mathcal{M}^\dagger \mathbf{b}_j + \left( \mathbf{b}_i^\dagger \mathbf{b}_g \mathcal{M}^\dagger \mathbf{b}_j + \mathbf{b}_i^\dagger \mathcal{M} \mathbf{b}_g^\dagger \right) \left( \bar{n}^{-1} - \mathcal{M}^\dagger \mathbf{b}_g \right) + \mathbf{b}_i^\dagger \mathbf{b}_g \mathcal{M}^\dagger \mathcal{M} \mathbf{b}_g^\dagger \mathbf{b}_j}{\left( \bar{n}^{-1} - \mathcal{M}^\dagger \mathbf{b}_g \right)^2 - \mathbf{b}_g^\dagger \mathbf{b}_g \mathcal{M}^\dagger \mathcal{M}} \bar{n} P. \quad (\text{B5})$$

In the continuous limit ( $N \rightarrow \infty$ ), we can exchange the vector products by integrals over the mass function and set  $\bar{n}_{\text{tot}} = \bar{n} N$ . This finally yields

$$\Sigma_{ij} = \langle \mathbf{b}_i \mathbf{b}_j \rangle \bar{n}_{\text{tot}} P + \frac{\langle \mathbf{b}_g^2 \rangle \langle \mathcal{M} \mathbf{b}_i \rangle \langle \mathcal{M} \mathbf{b}_j \rangle + (\langle \mathbf{b}_i \mathbf{b}_g \rangle \langle \mathcal{M} \mathbf{b}_j \rangle + \langle \mathcal{M} \mathbf{b}_i \rangle \langle \mathbf{b}_j \mathbf{b}_g \rangle) (\bar{n}_{\text{tot}}^{-1} - \langle \mathcal{M} \mathbf{b}_g \rangle) + \langle \mathbf{b}_i \mathbf{b}_g \rangle \langle \mathbf{b}_j \mathbf{b}_g \rangle \langle \mathcal{M}^2 \rangle}{(\bar{n}_{\text{tot}}^{-1} - \langle \mathcal{M} \mathbf{b}_g \rangle)^2 - \langle \mathbf{b}_g^2 \rangle \langle \mathcal{M}^2 \rangle} \bar{n}_{\text{tot}} P. \quad (\text{B6})$$

- 
- [1] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, et al., *ApJ* **517**, 565 (1999), arXiv:astro-ph/9812133.
  - [2] *BigBOSS*, URL <http://bigboss.lbl.gov>.
  - [3] *EUCLID*, URL <http://sci.esa.int/euclid>.
  - [4] *SDSS-III*, URL <http://www.sdss3.org>.
  - [5] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, *Phys. Rep.* **367**, 1 (2002), arXiv:astro-ph/0112551.
  - [6] M. Crocce and R. Scoccimarro, *Phys. Rev. D* **73**, 063519 (2006), arXiv:astro-ph/0509418.
  - [7] A. Taruya and T. Hiramatsu, *ApJ* **674**, 617 (2008), 0708.1367.
  - [8] T. Matsubara, *Phys. Rev. D* **78**, 083519 (2008), 0807.1733.
  - [9] K. Chuen Chan and R. Scoccimarro, *ArXiv e-prints* (2012), 1204.5770.
  - [10] T. Baldauf, U. Seljak, V. Desjacques, and P. McDonald, *ArXiv e-prints* (2012), 1201.4827.
  - [11] U. Seljak, *Physical Review Letters* **102**, 021302 (2009), 0807.1770.
  - [12] P. McDonald and U. Seljak, *J. Cosmology Astropart. Phys.* **10**, 7 (2009), 0810.0323.
  - [13] L. Guzzo, M. Pierleoni, B. Meneux, E. Branchini, O. Le Fèvre, C. Marinoni, B. Garilli, J. Blaizot, G. De Lucia, A. Pollo, et al., *Nature* **451**, 541 (2008), 0802.1944.
  - [14] W. J. Percival and M. White, *MNRAS* **393**, 297 (2009), 0808.0003.
  - [15] A. Cabré and E. Gaztañaga, *MNRAS* **393**, 1183 (2009), 0807.2460.
  - [16] M. Crocce, E. Gaztañaga, A. Cabré, A. Carnero, and E. Sánchez, *MNRAS* **417**, 2577 (2011), 1104.5236.
  - [17] C. Blake, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom, T. Davis, M. J. Drinkwater, K. Forster, D. Gilbank, et al., *MNRAS* **415**, 2876 (2011), 1104.2948.
  - [18] A. J. Ross, W. J. Percival, M. Crocce, A. Cabré, and E. Gaztañaga, *MNRAS* **415**, 2193 (2011), 1102.0968.
  - [19] A. Nusser, E. Branchini, and M. Davis, *ApJ* **744**, 193 (2012), 1106.6145.
  - [20] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, G. B. Poole, L. Campbell, Q. Parker, W. Saunders, and F. Watson, *MNRAS* **423**, 3430 (2012), 1204.4725.
  - [21] B. A. Reid, L. Samushia, M. White, W. J. Percival, M. Manera, N. Padmanabhan, A. J. Ross, A. G. Sánchez, S. Bailey, D. Bizyaev, et al., *ArXiv e-prints* (2012), 1203.6641.
  - [22] L. Samushia, B. A. Reid, M. White, W. J. Percival, A. J. Cuesta, L. Lombriser, M. Manera, R. C. Nichol, D. P. Schneider, D. Bizyaev, et al., *ArXiv e-prints* (2012), 1206.5309.
  - [23] A. Raccanelli, D. Bertacca, D. Pietrobon, F. Schmidt, L. Samushia, N. Bartolo, O. Doré, S. Matarrese, and W. J. Percival, *ArXiv e-prints* (2012), 1207.0500.
  - [24] N. Kaiser, *MNRAS* **227**, 1 (1987).
  - [25] J. A. Peacock and S. J. Dodds, *MNRAS* **267**, 1020 (1994), arXiv:astro-ph/9311057.
  - [26] R. Scoccimarro, *Phys. Rev. D* **70**, 083007 (2004), arXiv:astro-ph/0407214.
  - [27] J. L. Tinker, D. H. Weinberg, and Z. Zheng, *MNRAS* **368**, 85 (2006), arXiv:astro-ph/0501029.
  - [28] B. A. Reid and M. White, *MNRAS* **417**, 1913 (2011), 1105.4165.
  - [29] E. Jennings, C. M. Baugh, and S. Pascoli, *MNRAS* **410**, 2081 (2011), 1003.4282.
  - [30] M. Sato and T. Matsubara, *Phys. Rev. D* **84**, 043501 (2011), 1105.5007.
  - [31] J. Kwan, G. F. Lewis, and E. V. Linder, *ApJ* **748**, 78 (2012), 1105.1194.
  - [32] L. Samushia, W. J. Percival, and A. Raccanelli, *MNRAS* **420**, 2102 (2012), 1102.1014.
  - [33] S. de la Torre and L. Guzzo, *ArXiv e-prints* (2012), 1202.5559.
  - [34] D. Bianchi, L. Guzzo, E. Branchini, E. Majerotto, S. de la Torre, F. Marulli, L. Moscardini, and R. E. Angulo, *ArXiv e-prints* (2012), 1203.1545.
  - [35] P. J. E. Peebles, *The large-scale structure of the universe* (1980).
  - [36] E. V. Linder and R. N. Cahn, *Astroparticle Physics* **28**, 481 (2007), arXiv:astro-ph/0701317.



- [37] T. Okumura and Y. P. Jing, *ApJ* **726**, 5 (2011), 1004.3548.
- [38] R. A. Fisher, *Roy. Stat. Soc.* **98**, 39 (1935).
- [39] M. Tegmark, A. N. Taylor, and A. F. Heavens, *ApJ* **480**, 22 (1997), arXiv:astro-ph/9603021.
- [40] A. Heavens, *ArXiv e-prints* (2009), 0906.0664.
- [41] J. Sherman and W. J. Morrison, *Ann. Math. Stat.* **21**, 124 (1950).
- [42] M. S. Bartlett, *Ann. Math. Stat.* **22**, pp. 107 (1951), ISSN 00034851.
- [43] U. Seljak, N. Hamaus, and V. Desjacques, *Phys. Rev. Lett.* **103**, 091303 (2009), 0904.2963.
- [44] V. Springel, *MNRAS* **364**, 1105 (2005), arXiv:astro-ph/0505010.
- [45] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, et al., *ApJS* **180**, 330 (2009), 0803.0547.
- [46] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White, *ApJ* **292**, 371 (1985).
- [47] R. W. Hockney and J. W. Eastwood, *Computer simulation using particles* (1988).
- [48] C. B. Markwardt, in *Astronomical Data Analysis Software and Systems XVIII*, edited by D. A. Bohlender, D. Durand, & P. Dowler (2009), vol. 411 of *Astronomical Society of the Pacific Conference Series*, p. 251, 0902.2850.
- [49] J. Moré, *The Levenberg-Marquardt algorithm: Implementation and theory*, vol. 630 of *Lecture Notes in Mathematics* (Springer Berlin / Heidelberg, 1978).
- [50] N. Hamaus, U. Seljak, V. Desjacques, R. E. Smith, and T. Baldauf, *Phys. Rev. D* **82**, 043515 (2010), 1004.5377.
- [51] R. E. Smith, R. Scoccimarro, and R. K. Sheth, *Phys. Rev. D* **75**, 063512 (2007), arXiv:astro-ph/0609547.
- [52] N. Hamaus, U. Seljak, and V. Desjacques, *Phys. Rev. D* **84**, 083509 (2011), 1104.2321.
- [53] J. N. Fry and E. Gaztañaga, *ApJ* **413**, 447 (1993), arXiv:astro-ph/9302009.
- [54] A. F. Heavens, S. Matarrese, and L. Verde, *MNRAS* **301**, 797 (1998), arXiv:astro-ph/9808016.
- [55] P. McDonald, *Phys. Rev. D* **74**, 103512 (2006), arXiv:astro-ph/0609413.
- [56] H. Gil-Marín, C. Wagner, L. Verde, R. Jimenez, and A. F. Heavens, *MNRAS* **407**, 772 (2010), 1003.3238.
- [57] M. White, Y.-S. Song, and W. J. Percival, *MNRAS* **397**, 1348 (2009), 0810.1518.
- [58] G. M. Bernstein and Y.-C. Cai, *MNRAS* **416**, 3009 (2011), 1104.3862.
- [59] T. Okumura, U. Seljak, and V. Desjacques, *ArXiv e-prints* (2012), 1206.4070.
- [60] T. Baldauf, R. E. Smith, U. Seljak, and R. Mandelbaum, *Phys. Rev. D* **81**, 063531 (2010), 0911.4973.
- [61] E. Gaztañaga, M. Eriksen, M. Crocce, F. J. Castander, P. Fosalba, P. Marti, R. Miquel, and A. Cabré, *MNRAS* **422**, 2904 (2012), 1109.4852.
- [62] Y.-C. Cai and G. Bernstein, *MNRAS* **422**, 1045 (2012), 1112.4478.
- [63] M. Cacciato, O. Lahav, F. C. van den Bosch, H. Hoekstra, and A. Dekel, *MNRAS* **426**, 566 (2012), 1203.2616.
- [64] P. Simon, *ArXiv e-prints* (2012), 1203.6205.
- [65] E. Jullo, J. Rhodes, A. Kiessling, J. E. Taylor, R. Massey, J. Berge, C. Schimd, J.-P. Kneib, and N. Scoville, *ApJ* **750**, 37 (2012), 1202.6491.
- [66] F. van den Bosch, S. More, M. Cacciato, H. Mo, and X. Yang, *ArXiv e-prints* (2012), 1206.6890.
- [67] W. Hu, *ApJ* **522**, L21 (1999).
- [68] U. Pen, *MNRAS* **350**, 1445 (2004), arXiv:astro-ph/0402008.
- [69] R. Massey, J. Rhodes, R. Ellis, N. Scoville, A. Leauthaud, A. Finoguenov, P. Capak, D. Bacon, H. Aussel, J.-P. Kneib, et al., *Nature* **445**, 286 (2007).
- [70] P. Simon, A. N. Taylor, and J. Hartlap, *MNRAS* **399**, 48 (2009), 0907.0016.
- [71] J. T. VanderPlas, A. J. Connolly, B. Jain, and M. Jarvis, *ApJ* **727**, 118 (2011), 1008.2396.
- [72] U. Seljak and P. McDonald, *J. Cosmology Astropart. Phys.* **11**, 39 (2011), 1109.1888.
- [73] S. W. Allen, A. E. Evrard, and A. B. Mantz, *ARA&A* **49**, 409 (2011), 1103.4829.
- [74] R. E. Angulo, V. Springel, S. D. M. White, A. Jenkins, C. M. Baugh, and C. S. Frenk, *ArXiv e-prints* (2012), 1203.3216.
- [75] Y. Noh and J. D. Cohn, *ArXiv e-prints* (2012), 1204.1577.
- [76] E. Rozo, E. S. Rykoff, A. Evrard, M. Becker, T. McKay, R. H. Wechsler, B. P. Koester, J. Hao, S. Hansen, E. Sheldon, et al., *ApJ* **699**, 768 (2009), 0809.2794.
- [77] A. V. Kravtsov, A. Vikhlinin, and D. Nagai, *ApJ* **650**, 128 (2006), arXiv:astro-ph/0603205.
- [78] *ACT*, URL <http://physics.princeton.edu/act>.
- [79] *DES*, URL <http://www.darkenergysurvey.org>.
- [80] *eROSITA*, URL <http://mpe.mpg.de/erosita>.
- [81] *LSST*, URL <http://www.lsst.org/lsst>.
- [82] *SPT*, URL <http://pole.uchicago.edu>.
- [83] *WFXT*, URL <http://wfxt.pha.jhu.edu>.
- [84] U. Seljak, *MNRAS* **318**, 203 (2000), arXiv:astro-ph/0001493.
- [85] R. A. Skibba, F. C. van den Bosch, X. Yang, S. More, H. Mo, and F. Fontanot, *MNRAS* **410**, 417 (2011), 1001.4533.
- [86] Z. Zheng, I. Zehavi, D. J. Eisenstein, D. H. Weinberg, and Y. P. Jing, *ApJ* **707**, 554 (2009), 0809.1868.
- [87] Z. Zheng, A. A. Berlind, D. H. Weinberg, A. J. Benson, C. M. Baugh, S. Cole, R. Davé, C. S. Frenk, N. Katz, and C. G. Lacey, *ApJ* **633**, 791 (2005), arXiv:astro-ph/0408564.
- [88] Y. Cai, G. Bernstein, and R. K. Sheth, *MNRAS* **412**, 995 (2011), 1007.3500.
- [89] R. Laureijs, J. Amiaux, S. Arduini, J. . Auguères, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvel, A. Ealet, et al., *ArXiv e-prints* (2011), 1110.3193.
- [90] E. Majerotto, L. Guzzo, L. Samushia, W. J. Percival, Y. Wang, S. de la Torre, B. Garilli, P. Franzetti, E. Rossetti, A. Cimatti, et al., *MNRAS* **424**, 1392 (2012), 1205.6215.
- [91] U. Seljak, *J. Cosmology Astropart. Phys.* **3**, 4 (2012), 1201.0594.
- [92] J. Yoo, N. Hamaus, U. Seljak, and M. Zaldarriaga, *Phys. Rev. D* **86**, 063514 (2012), 1206.5809.